Intelligent, Spoken Medical Records: A Prototype

Wayne H. Ward, M.A., Ph.D., Rickard R. Holtmeier, M.B.A.
Berdy Medical Systems, Inc.
Rochelle Park, New Jersey

The last major barrier to broad-based acceptance of Computer-Based Patient Records (CPR) is the reluctance of physicians to use standard data input mechanisms. The past promises of speech recognition conceptually overcame this reluctance. Unfortunately, commercial speech recognition is not suitable, primarily because it is speaker dependant, requires discontinuous speech, and/or does not understand the utterance context.

This demonstration will show a prototype of a system, SmartVoice, that is speaker independent, recognizes continuous speech and understands the context. As a prototype, SmartVoice is limited to use in the physical examination process of a clinical encounter.

TECHNOLOGY

SmartVoice extends the current academic state-of-the-art research systems in artificial intelligence, natural language and speech understanding. These systems are already considerably advanced over what is commercially available in the medical field.

SmartVoice extends speech technology developed at Carnegie Mellon University (CMU). As a beginning step, Berdy Medical Systems (BMS) acquired the rights to use CMU's Sphinx-II speech recognition system, the Phoenix spoken language understanding system and the MINDS-II dialogue system. We then significantly extended and modified these technologies.

Medical Vocabulary

Proper speech recognition and context understanding require collection of a vocabulary. The physical examination vocabulary was collected from BMS's SmartClinic CPR and from the analysis of written and dictated physical examinations. The vocabulary is not only used to recognize words, but to place them in proper context.

For example, SmartVoice understands the term "anicteric sclerae" belongs in the HEEN category of the physical examination. It understands that the utterance "all other systems negative" means to list pertinent negatives for any body systems not otherwise mentioned with a positive finding. Utterances are actually "gisted" into appropriate database elements. We will extend this physical examination prototype to include all aspects of the clinical encounter.

Target Product

The final stage of development will be to use SmartVoice as the primary input mechanism for the comprehensive SmartClinic CPR. These products will be merged into SmartMedicine. For example, a physician will be able to say "Computer, I did a full physical examination, but found only swelling in the left ear. It appears to be an acute infection of the pinna. I will treat with Otobiotic, or generic equivalent, four drops in left ear q.i.d. If there is no abatement within 24 hours, she should come back again. In any case, she should see me in two weeks."

SmartVoice will understand this discussion and perform the following functions: Check the prescription for interactions, print the prescription, update the medications archive, create the physical examination, record the new diagnosis on the problem list, create the progress note, schedule the next visit, and send diagnosis and treatment information to the billing system for reimbursement.

DEMONSTRATION

We will invite physicians to dictate a physical examination into SmartVoice. No advance training will be used. The physician will be encouraged to use the same utterances he would use to dictate or write the results of a physical examination. He will be encouraged to speak normally, without any artificial word gaps. Because this is a prototype system, out of context language will not be handled.

When he finishes speaking to the computer, it will print out a standard format physical examination. We will then ask that he review the report for accuracy. Note that SmartVoice will not attempt to record the utterances word for word. It will attempt to produce a physical examination that properly expressed the intended findings.

CONCLUSION

SmartVoice and its architecture seem to work quite well for the limited prototype using the physical examination. The remaining effort is primarily collecting utterances and building semantic models for other components of the clinical encounter and the associated CPR. When this work is complete, "gisted," conversational, spoken communication with a CPR will significantly enhance CPR acceptance and use by physicians.

0195—4210/96/$5.00